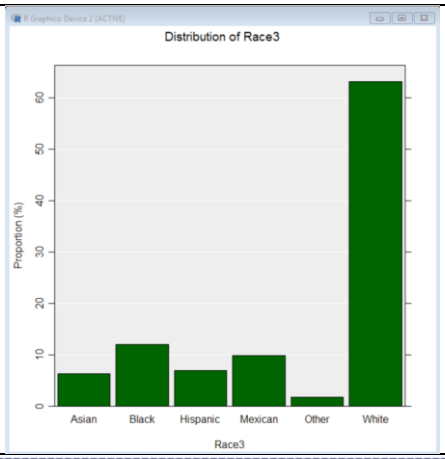
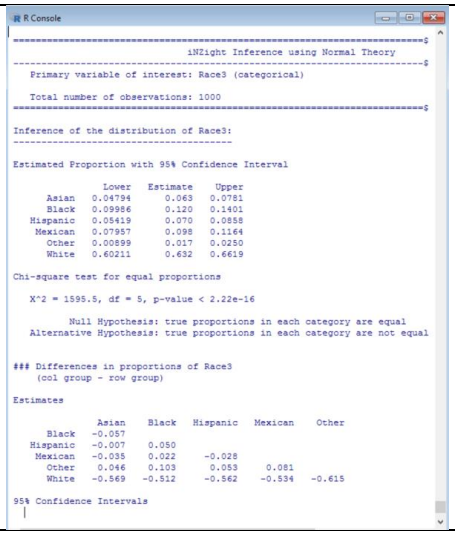
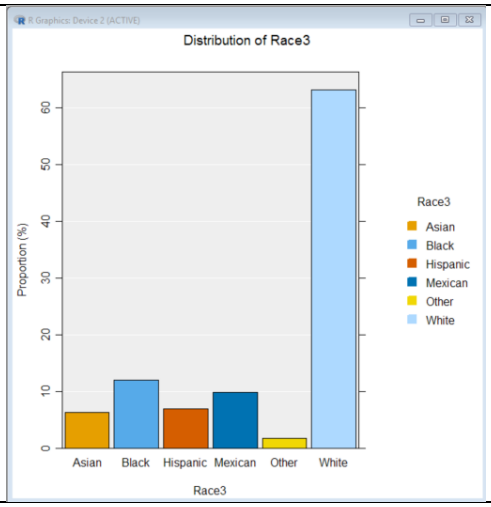
2.5 Exercise: Categorical variables – *R version*

Note: *Copying and pasting text (e.g. R code) from a pdf is not reliable. For that reason we have also provided this file in Word format (.docx) and also the code in a text file*

# R code	Output
# Setup library(iNZightPlots) library(FutureLearnData) data(nhanes_1000)	
# Plot the variable Race3 *# Because Race3 is **categorical** we get a **bar chart*** **iNZightPlot(Race3, data=nhanes_1000)**	
# Get a summary for a variable (*Race3*) *# Equiv. of Get Summary in iNZight* **getPlotSummary(Race3, data=nhanes_1000)**	iNZight Summary Primary variable of interest: Race3 (categorical) Total number of observations: 1000 Summary of the distribution of Race3: | | Asian | Black | Hispanic | Mexican | Other | White | Total | | --- | --- | --- | --- | --- | --- | --- | --- | | Count | 63 | 120 | 70 | 98 | 17 | 632 | 1000 | | Percent | 6.3% | 12.0% | 7.0% | 9.8% | 1.7% | 63.2% | 100% |
# Equivalent of *Get Inference* in iNZight getPlotSummary(Race3, data=nhanes_1000, **summary.type="inference", inference.type="conf"**)	

 # Colour by a variable (*Race3*) (default colour palette)

iNZightPlot(Race3, data=nhanes_1000, **colby=Race3)**



 # Create a new variable Race3.reord **to re-order** Race3
 # with the categories in frequency order

levels(nhanes_1000$**Race3)**

nhanes_1000**$Race3.reord** =
factor(nhanes_1000**$Race3, levels = c(**"White", "Black", "Mexican", "Hispanic", "Asian", "Other"**))**

iNZightPlot(**Race3.reord**, data=nhanes_1000)

COMMENTARY

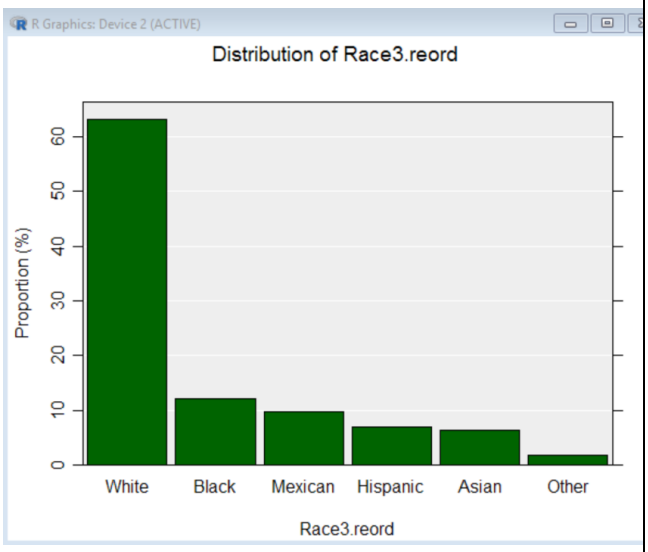
R *calls a* ***categorical*** *variable a "****factor****"*

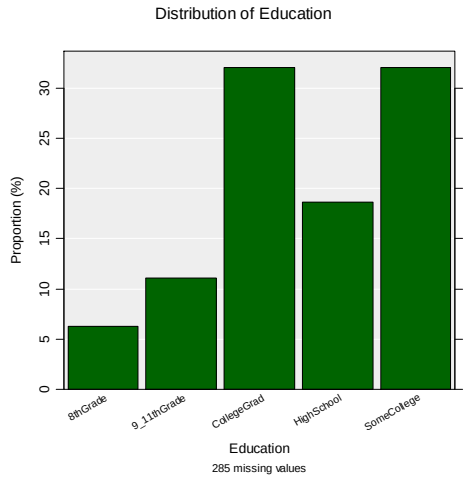
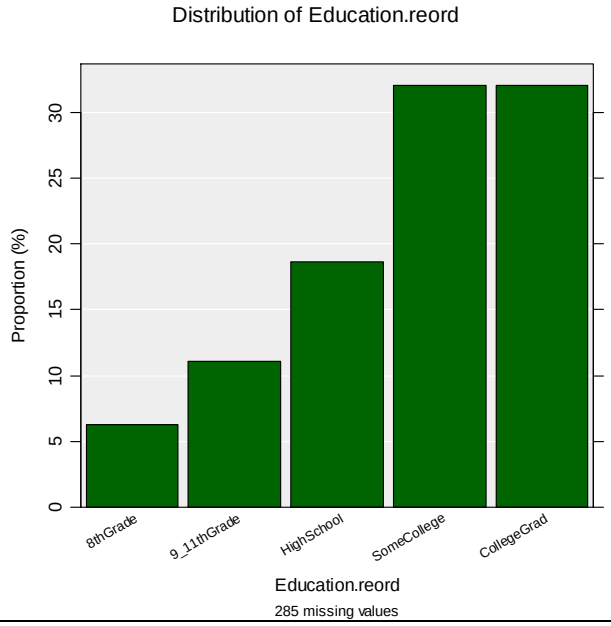
Show me the ***levels*** *of* ***Race3*** *(I can also see in the graph). Output is …*

[1] "Asian" "Black" "Hispanic" "Mexican" "Other" "White"

I can see *what the frequency order should be* ***from the graph.*** *(This can be done generally with code but the code is too complex to do at this stage)*

So I'll make ***Race3.reord*** *from* ***Race3*** *and put them in the order I want. (Getting the number of levels and spelling exactly right is crucial)*



# We'll do this again putting the levels of Education into a sensible order iNZightPlot(Education, data=nhanes_1000) **levels(**nhanes_1000$Education**)**	 [1] "8thGrade" "9_11thGrade" "CollegeGrad" "HighSchool" "SomeCollege"
# Create a new variable to re-order Education nhanes_1000$**Education.reord =** **factor(**nhanes_1000$**Education**, **levels =** **c(**"8thGrade","9_11thGrade","HighSchool", "SomeCollege","CollegeGrad"**))** levels(nhanes_1000$**Education.reord)**	[1] "8thGrade" "9_11thGrade" "HighSchool" "SomeCollege" "CollegeGrad"
iNZightPlot(**Education.reord**, data=nhanes_1000)	
iNZightPlot(**Education.reord**, data=nhanes_1000,**colby=Education.reord**) ***# Now change the colour palette** to rainbow colours* iNZightPlot(Education.reord, data=nhanes_1000,colby=Education.reord, **col.fun=rainbow**)	**COMMENTARY** ***Colour by** Education.reord* ***Col.fun** has to be a **colour palette function*** *There are lots of colour palette functions in R, many you have to install other packages to get. rainbow() is a generally available colour palette*

library(colorspace) iNZightPlot(Education.reord, data=nhanes_1000,colby=Education.reord, **col.fun=rainbow_hcl**)	*Using the **rainbow_hcl** colour function from the **colorspace** package*

- **Try repeating the above using other choices for variables and settings**

If you want to try installing some other R packages, in the R menus
Go **Packages** > **Install packages** . You will probably be asked to choose a CRAN mirror site.

Then you will be shown a list of packages to choose from.

Installing the package **viridis** and then loading it [via *library(viridis)*] will give you access to the colour functions: ***viridis***, ***magma***, and ***inferno***

Optional: *Try this new feature (interactive web graphics)*

We will export an iNZightPlot graph as an *Interactive HTML* file and open this file up in our default browser. If that is a modern browser like Chrome, Firefox or Safari (but not Internet Explorer) this will then give you an interactive version of the graph that lets you query it in various ways like hovering over bars or clicking them. Explore!

You can give such files to others. They do not need to be connected to iNZight to work.

Here is sample code:

```
# Make a plot and also store the output in myplot
myplot = iNZightPlot(Education.reord, data=nhanes_1000,colby=Education.reord)

# Specify a location to store an Interactive HTML file. I will call my file "myintplot.html"
# You will have to change the path to the file!
filepath = "C:/Users/myusername/Desktop/myintplot.html"

exportHTML(myplot, filepath)
browseURL(filepath)   #open the file up in my default browser
```

To discuss issues related to this Exercise,

go to **https://gitter.im/iNZightVIT/d2i-R-discussion**

To be able to post to the list you will have to set up a (free) account on ***Github***
https://github.com/login

If your question relates to an Exercise, say which one you are talking about!